

105

PATENT OFFICE
JAPANESE GOVERNMENT

This is to certify that the annexed is a true copy of the following application
as filed with this Office.

Date of Application : August 17, 1999

Application Number : P11-230216

Applicant(s) : NIPPON TELEGRAPH
AND TELEPHONE CORPORATION

June 2, 2000

Commissioner,
Patent Office Kouzou OIKAWA

Number of Certificate: 2000-3041005

N-105

日 本 国 特 許 庁
PATENT OFFICE
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日
Date of Application:

1999年 8月17日

出 願 番 号
Application Number:

平成11年特許願第230216号

出 願 人
Applicant(s):

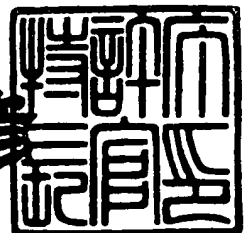
日本電信電話株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

2000年 6月 2日

特許庁長官
Commissioner,
Patent Office

近藤 隆彦



【書類名】 特許願

【整理番号】 NTTH115803

【提出日】 平成11年 8月17日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 15/40
G10L 1/02

【発明者】

【住所又は居所】 東京都千代田区大手町二丁目 3 番 1 号 日本電信電話株式会社内

【氏名】 大森 久美子

【特許出願人】

【識別番号】 000004226

【氏名又は名称】 日本電信電話株式会社

【代理人】

【識別番号】 100066153

【弁理士】

【氏名又は名称】 草野 卓

【選任した代理人】

【識別番号】 100100642

【弁理士】

【氏名又は名称】 稲垣 稔

【手数料の表示】

【予納台帳番号】 002897

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9806848

【プルーフの要否】 要

【書類名】明細書

【発明の名称】 検索データベースを記録した記録媒体並びに音声認識を用いたデータベース検索方法及び検索装置

【特許請求の範囲】

【請求項 1】 検索データベースは 2 階層構造であり、上位階層データはシステムが規定した実時間内に認識処理可能な対象データ数になるように構成し、

下位階層データは、検索キーを含むと共に、システムが規定した実時間内に認識処理が不可能なデータ数を含み、下位階層に含まれる各データは、必ず 1 つの上位階層中のデータに概念依存し、かつ 1 つの上位階層中の概念依存する下位階層のデータ数は前記実時間内に認識処理が可能なデータ数から構成し、

さらに、下位階層中の各データのアクセス頻度の偏りを示すアクセス頻度情報を持ち、下位階層データはアクセス頻度の高い方から順に、高頻度アクセスデータ群と、その他残りのデータを区別して保持すること

を特徴とする検索データベースを記録した記録媒体。

【請求項 2】 ユーザーが音声入力する検索キーに対して音声認識処理を行い、音声認識用検索データベースの中から、入力された検索キーを検索して同定することを目的とした音声認識を用いたデータベース検索方法において、

前記音声認識用検索データベースは、

2 階層構造であり、上位階層データは検索装置が規定した実時間内に処理可能な音声認識対象データ数になるように構成し、下位階層データは、ユーザーが音声入力する認識対象検索キーを含み、下位階層は前記実時間内に処理が不可能な音声認識対象データ数を含み、下位階層に含まれる各データは、必ず 1 つの上位階層中のデータに概念依存し、アクセス頻度に偏りを持ち、かつ前記実時間内に認識処理が可能なデータ数から構成し、

さらに、下位階層中の各データのアクセス頻度の偏りを示すアクセス頻度情報を持ち、下位階層データはアクセス頻度の高い方から順に、高頻度アクセスデータ群と、その他残りのデータを区別して保持し、

ユーザーから要求検索キーが音声入力されると、前記下位階層データの中の、高頻度アクセスデータ群に対して優先的に認識及び検索処理を行う手順 1 と、

検索キーに対する音声認識処理の結果算出した認識尤度の高い順に検索結果候補をユーザーに正誤性を問う確認質問をする手順 2 と、

手順 2 において、所定回数以下の確認質問で検索キーが同定できる場合は同定する手順 3 と、

手順 3 における所定回数の確認質問が、ユーザーから該当検索キーでないとすべて否定された場合には、ユーザーに対して、上位階層データを認識対象として、該当検索キーが依存する上位概念を尋ねる関連質問を行う手順 4 と、

関連質問に対するユーザーの応答を音声認識することにより、算出した上位概念候補の認識尤度を利用して、認識尤度の高い順に上位概念が同定できるまでユーザーの正誤性を問う確認質問を行う手順 5 と、

上位概念が同定できた時点で、同定した上位階層データに依存している下位階層データを認識対象データと限定して抽出する手順 6 と、

ユーザーが入力した検索キーに対して再び音声認識処理を行い、認識尤度の高い順にユーザーに対して正誤性を問う確認質問をすることで、音声検索キーの同定を行う手順 7 を備えたこと

を特徴とする音声認識を用いたデータベース検索方法。

【請求項 3】 音声入力部、認識対象データ抽出部、音声認識用検索データベースと応答データベースから構成される認識データベース、音声認識部、認識候補出力部及び音声出力部から音声認識を用いた検索装置を構成し、

ユーザーから要求検索キーが音声入力部に入力されると、

認識対象データ抽出部が音声認識検索データベースから認識対象データを抽出し、

この認識対象データを音声認識部で音声認識処理の結果算出した認識尤度の高い順に検索結果候補を認識候補出力部と音声出力部を介してユーザーに問い、ユーザーからの音声入力部からの応答を認識対象データ抽出部が応答データベースから認識対象データを抽出し、音声認識部で音声認識処理を行い、検索キーを同定する音声認識を用いた検索装置において、

前記音声認識用検索データベースが、2 階層構造であり、上位階層データは検索装置が規定した実時間内に認識処理可能な対象データ数になるように構成し、

下位階層データは、検索キーを含むと共に、前記実時間内に認識処理が不可能なデータ数を含み、下位階層に含まれる各データは、必ず 1 つの上位階層中のデータに概念依存し、かつ 1 つの上位階層中の概念依存する下位階層のデータ数は前記実時間内に認識処理が可能なデータ数から構成し、

さらに、下位階層中の各データのアクセス頻度の偏りを示すアクセス頻度情報を持ち、下位階層データはアクセス頻度の高い方から順に、高頻度アクセスデータ群と、その他残りのデータを区別して保持したこと

を特徴とする音声認識を用いた検索装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、各データがアクセス頻度に偏りを持ち、システムが規定した実時間内（すなわち、ユーザーにストレスを与えない時間内）に検索処理が不可能な対象データ数を有する大規模検索データベースを記録した記録媒体並びにユーザーが音声入力によって要求する検索キーの音声認識、検索及び同定を高速かつ正確に行う音声認識を用いたデータベース検索方法及び検索装置に関するものである。

【0002】

【従来の技術】

実時間内に処理不可能な大規模データベースに対する、ユーザーが要求する検索キーの同定を目的とした従来の検索方法は、実時間内に認識処理可能なデータ数に制限があることから、最初から要求検索キーを入力させるのではなく、認識対象を大規模データベース全体から実時間内に処理可能なデータ数に縮小できるような、検索キー候補の絞込みにつながる、検索補助キーの入力を促す、但し、検索補助キーの候補は、実時間内で認識処理可能なデータ数で構成され、ユーザーが要求する検索キーは必ず 1 つの検索補助キーを上位キーとして持ち、ユーザーにとって自分が検索を要求しようとしている検索キーの検索補助キー（上位キー）は単純明解なものであり、1 つの検索補助キーが有する下位キー（ユーザーが検索要求する検索キー）は実時間内に処理可能なデータ数で構成されるような

データを検索補助キーとして選択することにより検索キーの同定を行う。

【0003】

また、ユーザーが音声入力により要求する検索キーの同定を目的とした従来の検索方法では、まず始めに、検索補助キー（上位キー）に対して音声認識処理し、出力された検索補助キー（下位キー）候補を、正解という応答が得られるまで認識尤度の高い順にユーザーに提示する。検索補助キーが同定できたら、該当検索補助キーを上位キーとする検索キー（下位キー）候補を認識対象データとして抽出し、ユーザーに対して、本来検索要求したかった検索キー（下位キー）の入力を促す。検索キーに対しても検索補助キー同様に、音声認識処理の結果出力された候補を、認識尤度の高い順に正解という応答が得られるまでユーザーに提示することで、検索キーの同定を行う。

【0004】

【発明が解決しようとする課題】

検索装置において、認識処理は実時間内に認識用データベースを照合して認識尤度を算出し、認識結果を出力することのできるデータ数に限界があり、認識対象データ数が増えれば増えるほど、認識時間を要する。特に、機械による音声認識処理は、実時間内に認識用データベースを照合して認識尤度を算出し、認識結果を出力することのできる語彙数に限界があり、認識対象語彙数が増えれば増えるほど認識時間を要する。大規模データベースを認識対象とし、入力インターフェースとして音声を採用しているシステムは、音声認識処理の間、ユーザーを待機させるのが現状である。しかし、ユーザーは対オペレータのような対応をシステムに求めることから待機時間はユーザーにとってストレスにつながる。

【0005】

さらに、機械による音声認識処理は精度が100%ではない。特に、不特定話者の音声、話速が一定でない発話に対する音声、雑音が多い環境下での発話は認識が困難であり、さらに認識対象語彙素数が増えれば増えるほど認識精度は悪くなることから、特に大規模データベースを認識対象とする音声入力システムでは、ユーザーが音声入力した語彙の認識尤度を1位に算出することは困難であるのが現状である。

【0006】

音声を入力インターフェースとする大規模データベースを認識及び検索対象とした検索システムは、ユーザーにオペレータと対話しているかのような錯覚を与えるくらいの迅速かつ正確な応答を目標とすることから、認識処理時間と認識精度を補うようなシステムにとって効果的に認識対象語彙数の絞込みにつながるような質問形式を採用せざるを得ない。そのために、ユーザーが本当に検索を要求している検索キーを最初から入力させずに、検索キーの絞込みにつながる前記検索キーの同定を最初に行う。しかし、ユーザーにとっては、最終的に検索要求したい検索キーではなく、検索補助キーの入力を最初に強制され検索補助キーが確定できてから、検索キーを入力することはすることは入力したい検索キーの入力までにユーザーにとっては余計なやりとり（システムにとっては必要不可欠なやりとり）を強制されているようでストレスにつながる。

【0007】

本発明は、大規模データベースを認識対象として、ユーザーが入力する検索キーの同定を目的とする検索において、大規模データベース中のデータのアクセス頻度の偏りを利用して、ユーザーが本当に検索を要求している検索キーを最初から入力してもユーザーに認識の間、待機時間を与えずかつ認識精度の欠陥を感じさせずに認識及び検索を実行し、大規模データベースの中から該当検索キーを同定する迅速性及び自然性を兼ね備えた検索に用いられる検索データベースを記録した記録媒体並びに音声認識を用いたデータベース検索方法及び検索装置である。

【0008】

【課題を解決するための手段】

請求項1記載の発明は、ユーザーが検索キーを入力することにより検索装置が検索キーを同定するために用いられる検索データベースを記録した記録媒体において、検索データベースは2階層構造であり、上位階層データはシステムが規定した実時間内に認識処理可能な対象データ数になるように構成し、下位階層データは、検索キーを含むと共に、前記実時間内に認識処理が不可能なデータ数を含み、下位階層に含まれる各データは、必ず1つの上位階層中のデータに概念依存

し、かつ 1 つの上位階層中の概念依存する下位階層のデータ数は前記実時間内に認識処理が可能なデータ数から構成し、さらに、下位階層中の各データのアクセス頻度の偏りを示すアクセス頻度情報を持ち、下位階層データはアクセス頻度の高い方から順に、高頻度アクセスデータ群と、その他残りのデータを区別して保持することを特徴とする検索データベースを記録した記録媒体である。

【0009】

請求項 2 記載の発明は、ユーザーが音声入力する検索キーに対して音声認識処理を行い、音声認識用検索データベースの中から、入力された検索キーを検索して同定することを目的とした音声認識を用いたデータベース検索方法において、前記音声認識用検索データベースは、2 階層構造であり、上位階層は検索装置が規定した実時間内に処理可能な音声認識対象データ数になるように構成し、下位階層データは、ユーザが音声入力する認識対象検索キーを含み、下位階層は前記実時間内に処理が不可能な音声認識対象データ数を含み、下位階層に含まれる各データは、必ず 1 つの上位階層中のデータに概念依存し、アクセス頻度に偏りを持ち、かつ前記実時間内に認識処理が可能なデータ数から構成し、さらに、下位階層中の各データのアクセス頻度の偏りを示すアクセス頻度情報を持ち、下位階層データはアクセス頻度の高い方から順に、高頻度アクセスデータ群と、その他残りのデータを区別して保持し、ユーザーから要求検索キーが音声入力されると、前記下位階層データの中の、高頻度アクセスデータ群に対して優先的に認識及び検索処理を行う手順 1 と、検索キーに対する音声認識処理の結果算出した認識尤度の高い順に検索結果候補をユーザーに正誤性を問う確認質問をする手順 2 と、手順 2 において、所定回数以下の確認質問で検索キーが同定できる場合は同定する手順 3 と、手順 3 における所定回数の確認質問が、ユーザーから該当検索キーでないとすべて否定された場合には、ユーザーに対して、上位階層データを認識対象として、該当検索キーが依存する上位概念を尋ねる関連質問を行う手順 4 と、関連質問に対するユーザーの応答を音声認識することにより、算出した上位概念候補の認識尤度を利用して、認識尤度の高い順に上位概念が同定できるまでユーザーに正誤性を問う確認質問を行う手順 5 と、上位概念が同定できた時点で、同定した上位階層データに依存している下位階層データを認識対象データと限

定して抽出する手順6と、ユーザーが入力した検索キーに対して再び音声認識処理を行い、認識尤度の高い順にユーザーに対して正誤性を問う確認質問をすることで、音声検索キーの同定を行う手順7を備えた音声認識を用いたデータベース検索方法である。

【0010】

また、請求項3記載の発明は、音声入力部、認識対象データ抽出部、音声認識用検索データベースと応答データベースから構成される認識データベース、音声認識部、認識候補出力部及び音声出力部から音声認識を用いた検索装置を構成し、ユーザーから要求検索キーが音声入力部に入力されると、認識対象データ抽出部が音声認識検索データベースから認識対象データを抽出し、この認識対象データを音声認識部で音声認識処理の結果算出した認識尤度の高い順に検索結果候補を認識候補出力部と音声出力部を介してユーザーに問い、ユーザーからの音声入力部からの応答を認識対象データ抽出部が応答データベースから認識対象データを抽出し、音声認識部で音声認識処理を行い、検索キーを同定する音声認識を用いた検索装置において、前記音声認識用検索データベースが、2階層構造であり、上位階層データはシステムが規定した実時間内に認識処理可能な対象データ数になるように構成し、下位階層データは、検索キーを含むと共に、前記実時間内に認識処理が不可能なデータ数を含み、下位階層に含まれる各データは、必ず1つの上位階層中のデータに概念依存し、かつ1つの上位階層中の概念依存する下位階層のデータ数は前記実時間内に認識処理が可能なデータ数から構成し、さらに、下位階層中の各データのアクセス頻度の偏りを示すアクセス頻度情報を持ち、下位階層データはアクセス頻度の高い方から順に、高頻度アクセスデータ群と、その他残りのデータを区別して保持した音声認識を用いた検索装置である。

【0011】

本発明は、高頻度アクセスデータ群に該当検索キーが含まれていれば、検索キーの補助となる上位概念を尋ねる関連質問をせずに、ユーザが本来検索要求する検索キーの入力のみで、かつ実時間内に検索キー同定が可能であり、該当検索キーが高頻度アクセスデータ群に含まれていない場合も、ユーザーにとっては従来と違ってシステムにとって効果的な絞込みを行うための上位概念を尋ねる補助的

な質問から強制的にスタートするのではなく、本来検索要求する検索キーを始めに入力し、次に補助的な上位概念の入力をするという自然さを兼ね備えた高速検索手法である。また、ユーザーに音声認識処理に要する時間及び認識精度の欠陥を感じさせることなく、かつユーザーが検索を要求している検索キーを最初に入力させるという自然なシステムとのやりとりのもとに、実時間内に処理不可能なアクセス頻度に偏りを有するデータから構成される大規模音声認識用検索データベースの中から、ユーザーが音声入力する検索キーを同定できる。

【0012】

【作用】

音声認識精度が、100%であると仮定した場合、実時間で認識処理を実現するために、ユーザーが本来検索要求したい検索キーからではなく、認識対象語彙数を縮小化するために上位概念を検索補助キーとして始めに尋ね、上位概念が同定できたら上位概念に依存している下位概念を検索対象として抽出して検索キーの入力を促す従来の手法は、実時間音声認識処理による候補確定に $T1$ （秒）かかるとすると、ユーザーとの間で上位概念（検索補助キー）と検索キーの2度の同定作業が行われることから $2 \times T1$ （秒）要することになる。一方、下位階層の高頻度アクセスデータ群を80%のアクセス頻度を有するデータで構成し、ユーザーが本来検索要求したい検索キーの入力を最初に促し、高頻度アクセスデータ群に対して優先的に検索処理を行う本発明方法は、高頻度アクセスデータ群に該当検索キーが含まれている場合は $T1$ （秒）、含まれていない場合は上位概念を最初に尋ね絞り込む上記手法をとることから $2 \times T1$ （秒）要することから、全体で $0.8 \times T1 + 0.2 \times 2 \times T1 = 1.2 \times T1$ （秒）要することになり、本発明手法の方が検索キー同定に要する期待値は小さくなる。

【0013】

実際に、音声認識精度は100%でないことから上記の処理時間で検索処理が終了することは困難であるが、音声認識装置が50%の確率で1位に正解候補を出力、40%の確率で2位に正解候補を出力、10%の確率で3位に正解候補を出力するという認識精度を有し、認識対象データベース中に正解が含まれている場合は、3位までに正解候補を出力すると仮定した場合、ユーザーに対して出力候補が正解か

どうか問うのに T_1 (秒) 要するとすると、従来の手法は、上位概念の同定に $0.5 \times T_1 + 0.4 \times 2 \times T_1 + 0.1 \times 3 \times T_1 = 1.6 \times T_1$ (秒) (正解が 2 位の場合、正誤性を 2 回問うことから、 $2 \times T_1$ (秒) が正誤確認時間となる) 要する。そして、上位概念を用いて認識対象を実時間処理可能な語彙数に絞り込み、ユーザの要求している検索キーの同定に $1.6 \times T_1$ (秒) 要することから合計で $1.6 \times T_1 + 1.6 \times T_1$ (秒) 要する。一方、本発明手法は、同様の音声認識精度の下で、高頻度アクセスデータ群を 80% のアクセス頻度を有するデータで構成する場合、高頻度アクセスデータ群に該当検索キーが含まれている場合、下位階層から得られる検索キー候補の正誤を問う確認質問を 2 回までとすると、2 位までに正解検索キー候補が出力される場合は検索キー同定に $0.8 \times 0.5 \times T_1 + 0.8 \times 0.4 \times 2 \times T_1 = 1.04 \times T_1$ (秒) 要し、高頻度アクセスデータ群に該当検索キーが含まれていても上位 2 位までに正解検索キーが出力されない場合は上位概念を尋ね、検索範囲を絞込む従来の手法をとることから $0.5 \times T_1 + 0.4 \times 2 \times T_1 + 0.1 \times 3 \times T_1 = 1.04 \times T_1$ (秒) 要するのが全体の 10% (3 位に正解候補が出力される確率) なので、 $1.6 \times T_1 \times 0.1 = 0.16 \times T_1$ (秒) 要する。また該当検索キーが高頻度アクセスデータ群に含まれていない場合も従来の手法をとることから $1.6 \times T_1$ (秒) 要し、これが全体の 20% (アクセス頻度が 20% に満たない単語が該当検索キーである場合) なので $1.6 \times T_1 \times 0.2 = 0.32 \times T_1$ (秒) 要する。従って音声認識精度が 100% ではない場合、本発明の検索手法は合計で、 $1.04 \times T_1 + 0.16 \times T_1 + 0.32 \times T_1 + 1.52 \times T_1 = 1.52 \times T_1$ (秒) 要することになり、本発明による検索方法の方が従来方法と比較して検索キーの同定に要する時間の期待値が半分以下に押さえられ、さらにシステムにとって効率の良い絞り込みにつながる補助的な質問から始まるのではなく、ユーザが本来要求している検索キーを最初に入力するという自然性も兼ね備えている。

【0 0 1 4】

【発明の実施の形態】

図 1 は本発明による検索装置の実施形態の一例を示すブロック図である。

この検索装置 1 は、音声入力部 2、認識対象データ抽出部 3、音声認識部 4、認識候補出力部 5、音声出力部 6 から構成される。

認識対象データ抽出部 3 において、音声認識用検索データベース 7-1、応答

データベース 7-2 からなる認識データベース 7 を使用する。音声認識部 4 では音声認識装置 8 を、音声出力部 6 では音声出力装置 9 を使用する。

【0015】

図 2 に本発明による記録媒体に記録する音声認識用検索データベース 7-1 の全体像の一例を示す。

音声認識用検索データベース 7-1 は、上位概念と下位概念の 2 階層で構成し、ユーザーが検索を要求する検索キーは下位階層データである。上位階層は実時間内に処理可能な語彙数を有し、下位階層は実時間内に処理不可能な大規模語彙数を有する。下位階層データは必ず 1 つの上位階層のデータと依存関係を持ち、1 つの上位概念に依存する下位階層データ数は実時間内に処理できる数で構成される。また、下位階層の大規模データに対してアクセス頻度の偏りを利用して、アクセス頻度の高い順に上から、実時間内に処理できるデータ数だけ「頻」マーク付けしてその他の「稀」とマーク付けされたデータとは別メモリに保存することで、高頻度アクセスデータ群を作成する。

【0016】

検索装置 1 は、音声入力部 2 でユーザーから入力された音声が入力されると、入力された音声に応じて認識対象とするデータの識別を認識対象データ抽出部 3 で行う。

本発明の手順を図 8 に示す。

(S1) ユーザーから音声入力部 2 に対して検索キーが入力されると、(S2) 認識対象データ抽出部 3 では、まず始めに優先的に認識及び検索を行う音声認識用データベース 7-1 の下位階層データのうち、高頻度アクセスデータ群を認識対象データに指定する。

【0017】

(S3) 音声認識部 4 で音声認識処理した結果が (S4) 認識候補出力部 5 にて出力される。この際、算出された認識尤度の高い順に認識候補は出力される。(S5) 音声出力部 6 は、認識尤度の高い順に出力された検索キー候補をユーザーに提示し正誤を問う。(S8) 但し、正誤を問う回数はあらかじめ検索装置 1 によって規定されている。(S6) 音声入力部 2 から正誤を問う確認質問に対する応答が入力されると

、認識対象データ抽出部 3 は、認識データベース 7 の応答データベース 7-2 を認識対象データに指定し、(S7)「はい」という応答が音声認識部 4、認識候補出力部 4 にて確認できたら、音声出力部 6 にて検索キーが同定できたことをユーザーに知らせる。(S9)所定回数の検索キー候補の確認質問がすべてユーザーから否定された（音声認識部 4、認識候補出力部 5 にて「いいえ」という応答を出力）場合、音声出力部 6 は、上位階層データに含まれる検索キーの上位概念を尋ねる関連質問をユーザーに対して行う。(S10)関連質問に対する応答が、音声入力部 2 により入力されると、認識対象データ抽出部 3、音声認識部 4 は、認識した上位概念に依存している下位階層データを音声認識用検索データベース 7-1 から認識対象として抽出し、ユーザーが最初に入力した検索キーを再び音声認識部 4 にて認識する。(S11)認識候補出力部 5 にて認識尤度の高い順に出力された検索キー候補の正誤を尋ねる確認質問を音声出力部 6 から出力する。(S12)確認質問に対してユーザーから「はい」という応答が得られるまで確認を繰り返す。(S13)「はい」という応答が確認できたら検索キーが同定できたことをユーザーに知らせる。

【0 0 1 8】

【実施例】

図 1 を用いて本発明の検索方法により、日本全国の市区町村の同定を行う流れを説明する。

市区町村同定において、高頻度アクセスデータ群を認識対象とした場合に出力された検索キー候補に対して、認識尤度の高い順にユーザーに正誤を確認する確認質問は 3 回と規定する。市区町村同定の際の、音声認識用検索データベースの一例を図 3 に示す。検索キーとなる市区町村は音声認識用検索データベースの下位階層データであり、上位階層データとして日本全国の都道府県を選択する。都道府県は 47 という実時間内に処理可能な語彙数からなり、下位階層の市区町村は必ず属する都道府県を有し、1 つの都道府県に属する市区町村は最も市区町村を多く有する場合で 50 であることから、実時間内に処理可能であると考えられる。また、市区町村に対するアクセス頻度は、電話番号案内などに市区町村同定が使用される場合のアクセス頻度を利用して、アクセス頻度の高い上位 50（実

時間内に処理可能な語彙数)の市区町村を高頻度アクセスデータ群に指定する。

図4に高頻度アクセスデータ群に属する市区町村の一例を示す。

【0019】

始めに、「横浜市」を同定する場合について説明する。

音声入力部2から「横浜市」が入力されると、認識対象データ抽出部3は、音声認識用検索データベース7-1の下位階層データのうち高頻度アクセスデータ群に属する市区町村を認識対象データとして抽出する(例えば、図3において、札幌市、函館市、中央区、鹿児島市・・・)。「横浜市」は高頻度のアクセスデータ群に含まれているデータである。音声認識部4にて音声認識した結果が認識候補出力部5にて認識尤度の高い順に出力される。図5に出力結果の一例(1位横須賀市、2位横浜市、3位四つ木市・・・)を示す。

【0020】

音声出力部6は認識尤度の高い順にユーザーに対して検索キー候補の正誤を問う。

図5より正解の横浜市は2位に出力されていることから、確認質問を2回繰り返すことにより横浜市が同定できる。

次に、「横川市」を同定する場合について説明する。

【0021】

「横川市」は高頻度アクセスデータ群に含まれていないデータである。音声入力部2から「横川市」が入力されると、認識対象データ出力部3は高頻度アクセスデータ群を認識対象として抽出し、音声認識部4にて音声認識する。図6に認識候補出力部5から出力された結果の一例を示す。図6に示した結果より、「横浜市」「横須賀市」「横山市」の順にユーザーに対して正誤を問う確認質問を音声出力部6から出力する。音声入力部2から入力された確認質問の応答がすべて「いいえ」であることから、検索装置はユーザーに対して、検索キーである「横川市」の所属する都道府県の入力を音声出力部6より促す。ユーザーからの応答「群馬県」が音声入力部2から入力されると、認識対象データ抽出部3は群馬県に属する下位階層データ、すなわち群馬県内の41市区町村を認識対象データとして抽出する。そして改めて、音声認識部3にて「横川市」の認識を行い、認識

広報出力部 5 にて検索キー候補を出力する。図 7 に出力結果の一例を示す。音声出力部 6 にて認識尤度の高い順にユーザーに正誤を問う確認質問を行う。「横川市」は認識候補の 1 位であることから確認質問を 1 度することで同定される。

【0 0 2 2】

【発明の効果】

以上の説明からわかるように、音声認識対象語彙数が大規模な場合、現在の音声認識技術では実時間に処理可能な語彙数には限界があり、語彙数が多ければ多いほど認識精度が落ちることから、システムはユーザーの検索結果要求検索キーではなく、システムにとって実時間に精度よく認識できるような、小規模な検索対象語彙数に絞り込むことができるような効率の良い検索補助キーの入力を最初に強制する。

【0 0 2 3】

この発明によれば、認識要求検索データベースを 2 階層で構成し、ユーザーが検索を要求する検索キーを下位階層、上位には下位階層と依存関係にあるような実時間で処理可能な語彙数からなる検索補助キーを含むように、さらに 1 つの上位階層中のデータに依存する下位階層の検索キー候補は実時間処理可能な語彙数で構成されるよう上位階層を構成するデータを選択し、下位階層データの中でアクセス頻度の高い上位の実時間処理可能な語彙数は別のメモリに保存しておくことで、高頻度アクセスデータ群として優先的に検索及び認識対象とする。本発明はこのような工夫のもので、高頻度アクセスデータ群に該当検索キーが含まれていれば、検索キーの補助となる上位概念を尋ねる関連質問をせずに、ユーザーが本来検索要求する検索キーの入力のみで、かつ実時間内に検索キー同定が可能であり、該当検索キーが高頻度アクセスデータ群に含まれていない場合も、ユーザーにとっては従来と違ってシステムにとって効果的な絞り込みを行うための上位概念を尋ねる補助的な質問から強制的にスタートするのではなく、本来検索要求する検索キーを始めに入力し、次に補助的な上位概念の入力をするという自然さを兼ね備えている高速検索が可能な検索方法である。

【図面の簡単な説明】

【図 1】

本発明による検索装置のブロック図。

【図 2】

音声認識用検索データベースの構成図。

【図 3】

市区町村同定例の音声認識用検索データベースの一例を示す図。

【図 4】

市区町村同定例の高頻度アクセスデータ群の一例を示す図。

【図 5】

「横浜市」同定の場合の音声検索キー認識結果を示す図。

【図 6】

高頻度アクセスデータ群を認識対象とした場合の「横川市」同定の場合の音声検索キー認識結果を示す図。

【図 7】

群馬県内の市区町村を認識対象とした場合の「横川市」同定の場合の音声検索キー認識結果を示す図。

【図 8】

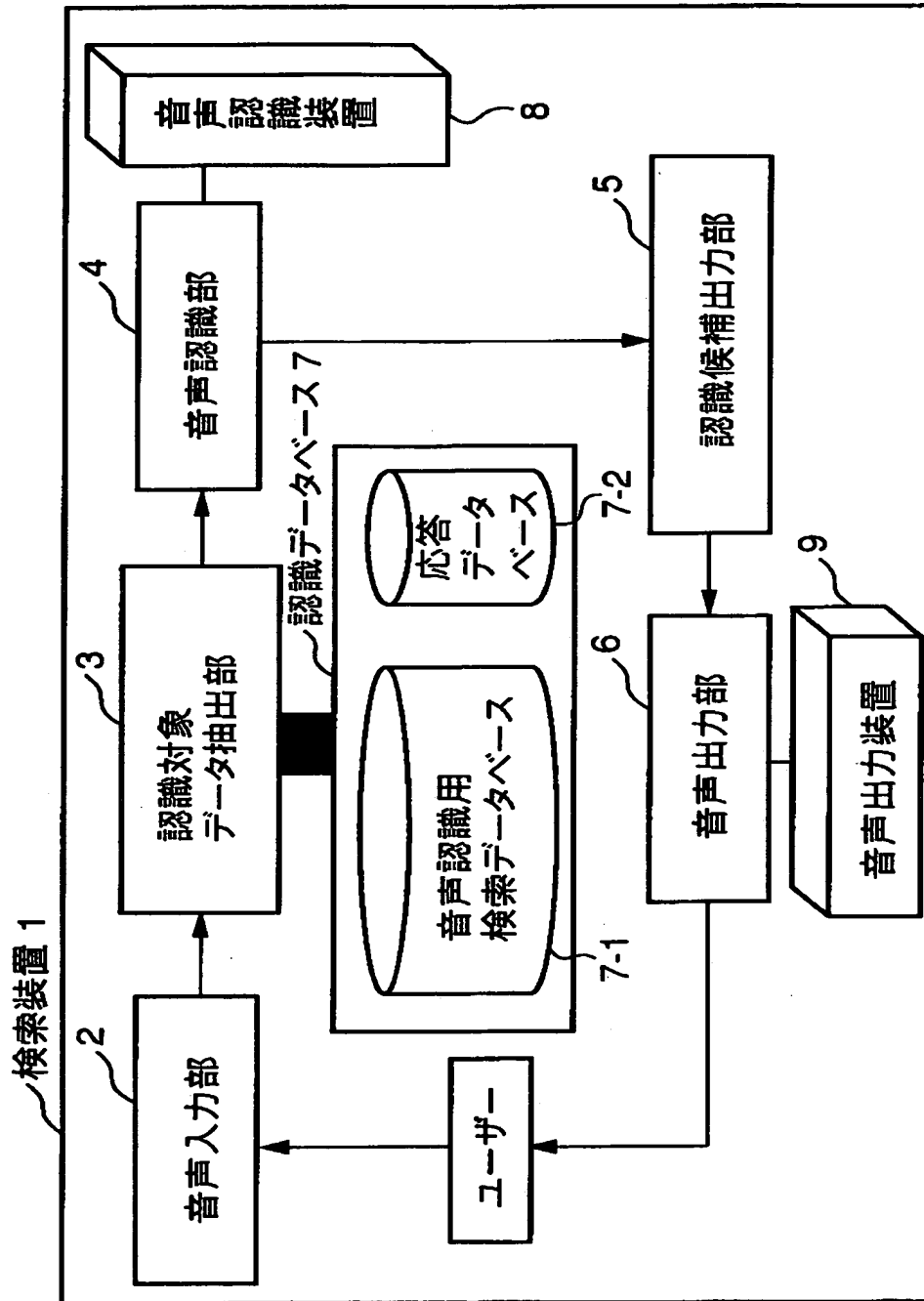
本発明の手順を示すフローチャート。

【符号の説明】

- 1 検索装置
- 2 音声入力部
- 3 認識対象データ抽出部
- 4 音声認識部
- 5 認識候補出力部
- 6 音声出力部
- 7 認識データベース
- 7-1 音声認識用検索データベース
- 7-2 応答データベース
- 8 音声認識装置
- 9 音声出力装置

【書類名】 図面

【図 1】



本発明による検索装置のブロック図

図 1

【図 2】

音声認識用検索データベース7-1

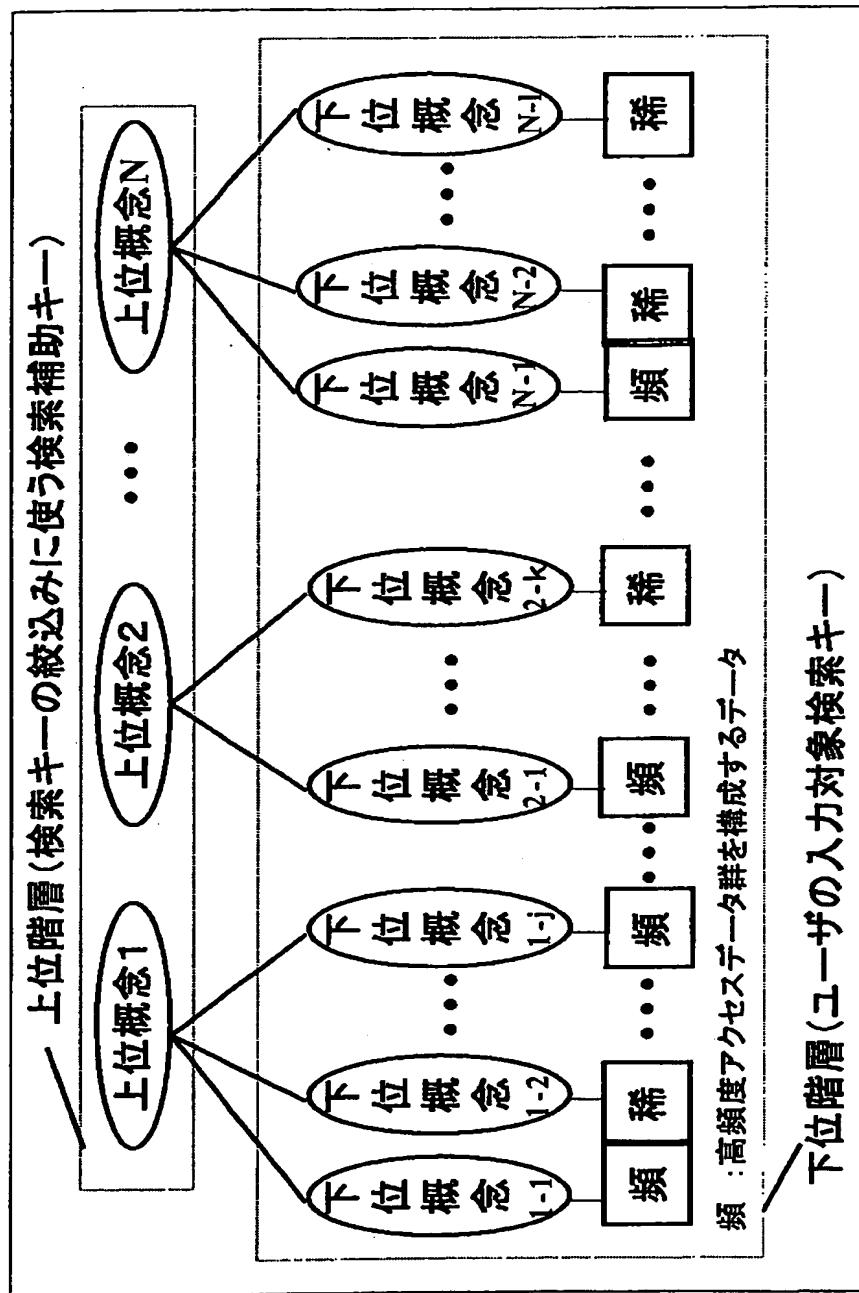


図 2: 音声認識用検索データベースの構成図

【図 3】

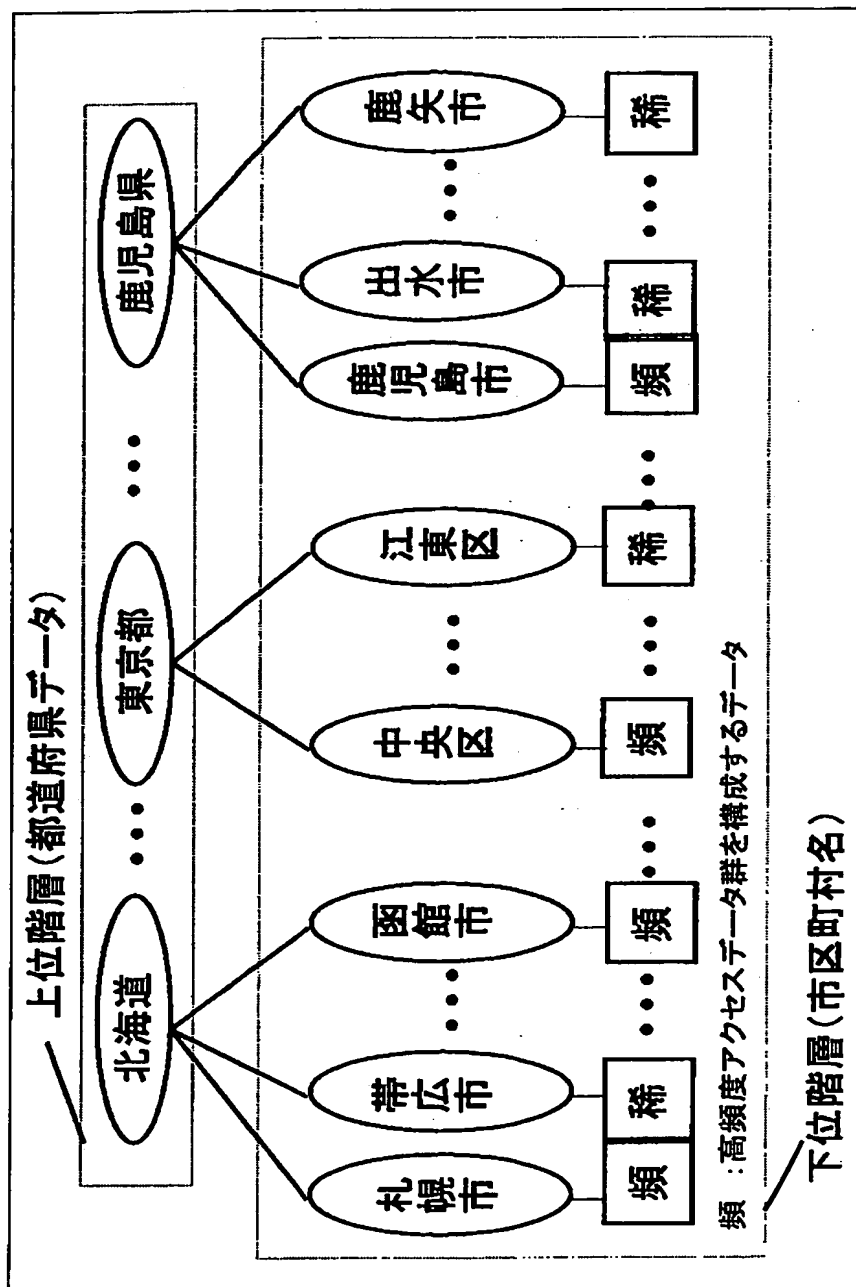


図 3: 市区町村同定例の音声認識用検索データベースの一例

【図 4】

アクセス頻度	市区町村
1位	港区
2位	中央区
3位	渋谷区
4位	横浜市
5位	新宿区
・	・
・	・
49位	大阪府
50位	舞浜市

アクセス頻度の高い順

図 4：市区町村同定例の高頻度アクセスデータ群の一例

【図 5】

順位	市区町村
1位	横須賀市
2位	横浜市
3位	四つ木市
4位	舞浜市
・	・
・	・
50位	港区

認識尤度の高い順 →

図 5: 「横浜市」同定の場合の音声検索キー認識結果

【図 6】

順位	市区町村
1位	横浜市
2位	横須賀市
3位	横山市
4位	美浜市
・	・
・	・
50位	千葉市

認識尤度の高い順 →

図 6: 「横川市」同定の場合の音声検索キー認識結果
(高頻度アクセスデータ群を認識対象とした場合)

【図 7】

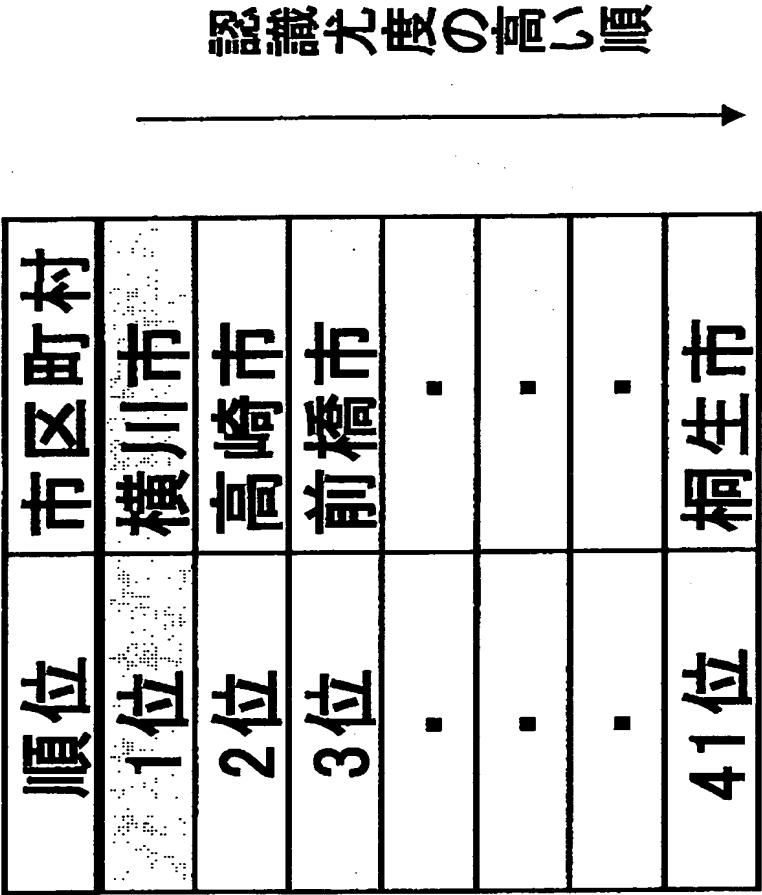


図 7: 「横川市」同定の場合の音声検索キー認識結果
(群馬県内の市区町村を認識対象とした場合)

【図 8】

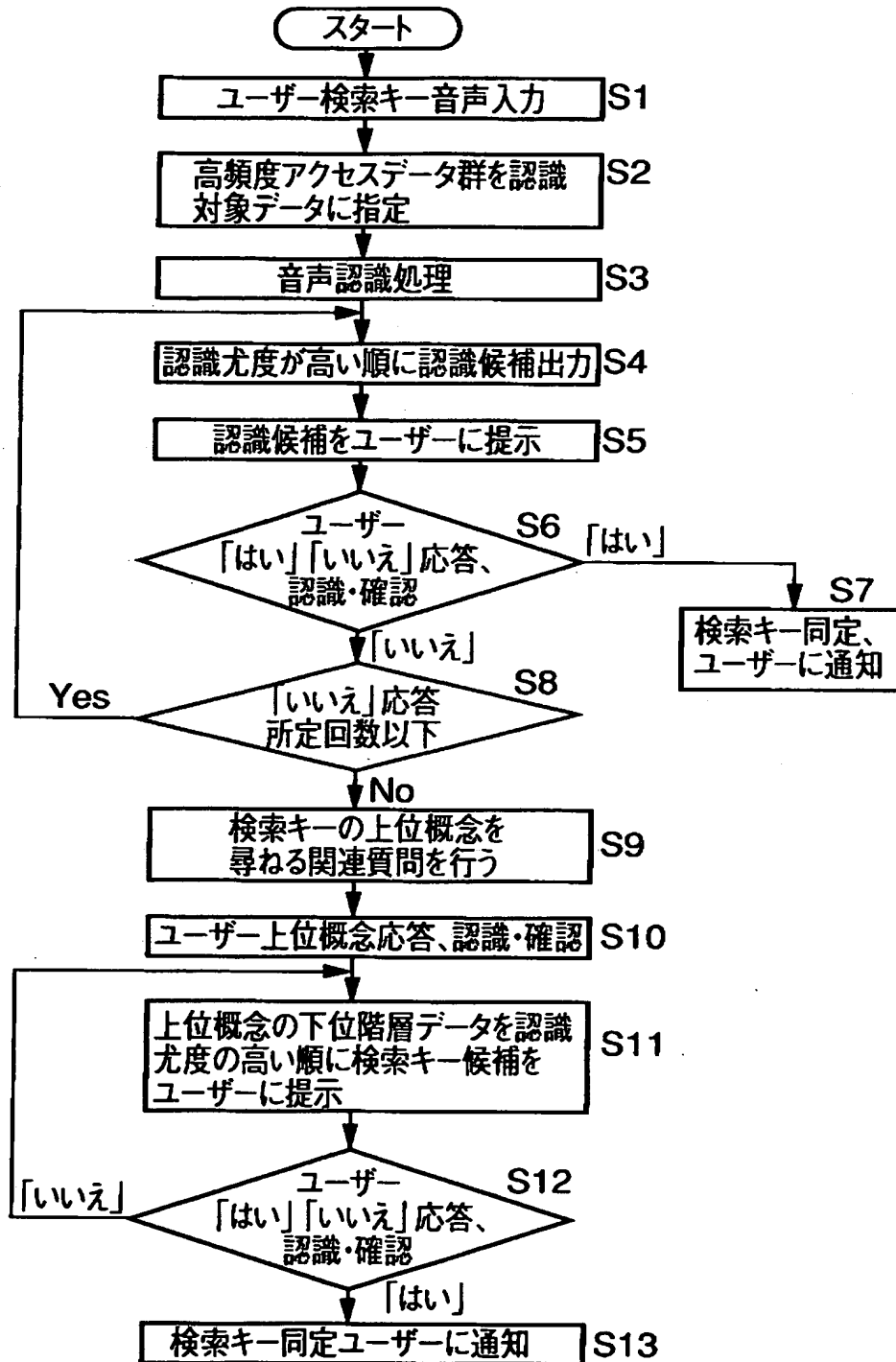


図 8

【書類名】 要約書

【要約】

【課題】 ユーザーが入力する検索キーを大規模データベースを認識対象とし、高速で、かつ実時間内認識処理が可能な検索を提供する。

【解決手段】 検索データベースが、2階層構造であり、上位階層データはシステムが規定した実時間内に認識処理可能な対象データ数になるように構成し、下位階層データは、検索キーを含むと共に、システムが規定された実時間内に認識処理が不可能なデータ数を含み、下位階層に含まれる各データは、必ず1つの上位階層中のデータに概念依存し、かつ1つの上位階層中の概念依存する下位階層のデータ数は前記実時間内に認識処理が可能なデータ数から構成し、さらに、下位階層中の各データのアクセス頻度の偏りを示すアクセス頻度情報を持ち、下位階層データはアクセス頻度の高い方から順に、高頻度アクセスデータ群と、その他残りのデータを区別して保持する。

【選択図】 図2

出 願 人 履 歴 情 報

識別番号 [000004226]

1. 変更年月日	1999年 7月15日
[変更理由]	住所変更
住 所	東京都千代田区大手町二丁目3番1号
氏 名	日本電信電話株式会社